

# Group 7

Body Fat Prediction

# Contents

1. Introduction
2. Model Selection and Results
3. Final Model Interpretation
4. Shiny App

# Introduction

Background and Data Cleaning

# Data cleaning

- **Scale**

We changed the units of weight and height to kg and cm because all of the other variables in the dataset are measured in cm.

- **Height**

We fixed a person who is only 74.94 cm tall using the relationship between bmi, weight and height ( $\text{bmi} = \text{kg}/\text{m}^2$ ).

- **Bodyfat**

We removed the person with 0% body fat percentage, because this is scientifically impossible and we couldn't manage to fix this.

# Model Selection and Results

Model Selection, Visualization, Result Discussion

# Models Selection

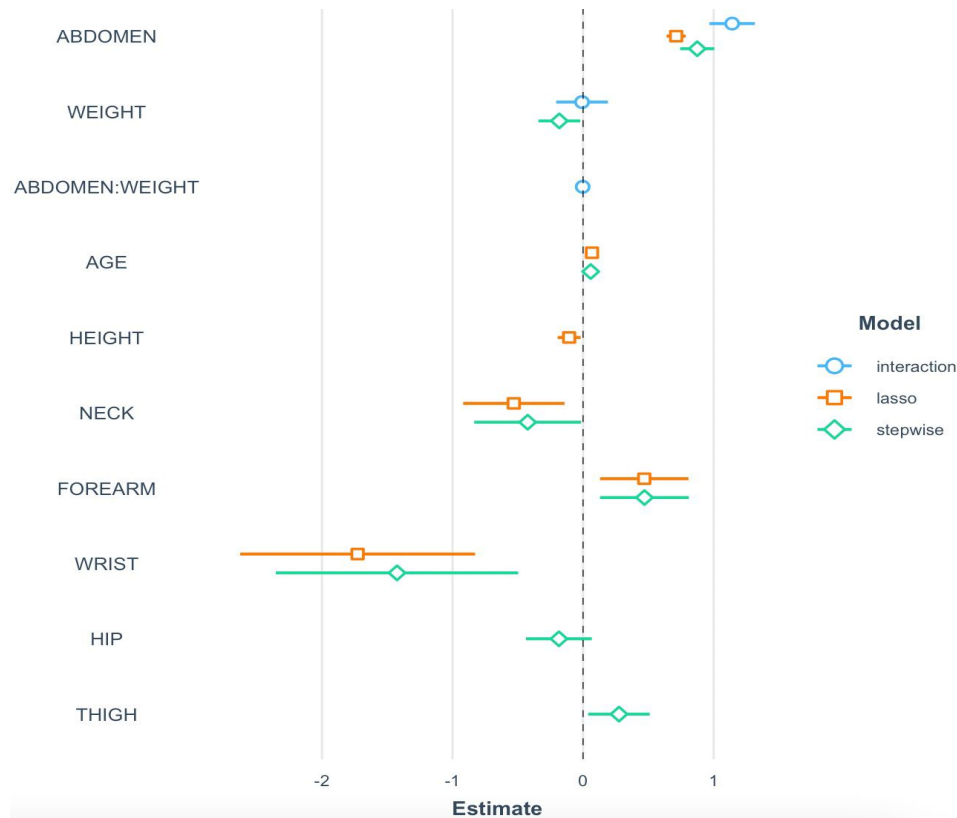
- **Stepwise linear regression model without interaction**  
Weight, age, neck, abdomen, thigh, forearm, wrist are significant
- **Lasso regression**  
Age, height, neck, abdomen, forearm, wrist are significant

**There are too many predictors and we want a simpler model**

1. The two models above show that weight, age, abdomen, height are four predictors significant and easy to measure
  2. We fit a linear model only consider these four predictors, and find that only abdomen and weight are significant
  3. We suggest that there are interaction effect between abdomen and weight
  4. So we fit a linear model with interaction
- **Linear regression with interaction**  
Abdomen and the interaction between abdomen and weight are significant

# Visualization of Results

- This plot shows the 95% CI of all the significant predictors in our three models
- We can see that estimates for each predictor are close to each other



# Discussions of results

- MSE and  $R^2$  are similar in these three models
- Overall p-value are all smaller than 0.05
- Among them, the interaction model uses least amount of predictors and is the simplest one

Model	MSE	$R^2$	Overall p-value
stepwise	15.12	0.74	$<2e-16$
Interaction model	16.08	0.72	$<2e-16$
lasso	15.67	0.73	$<2.2e-16$



# Final Model Interpretation

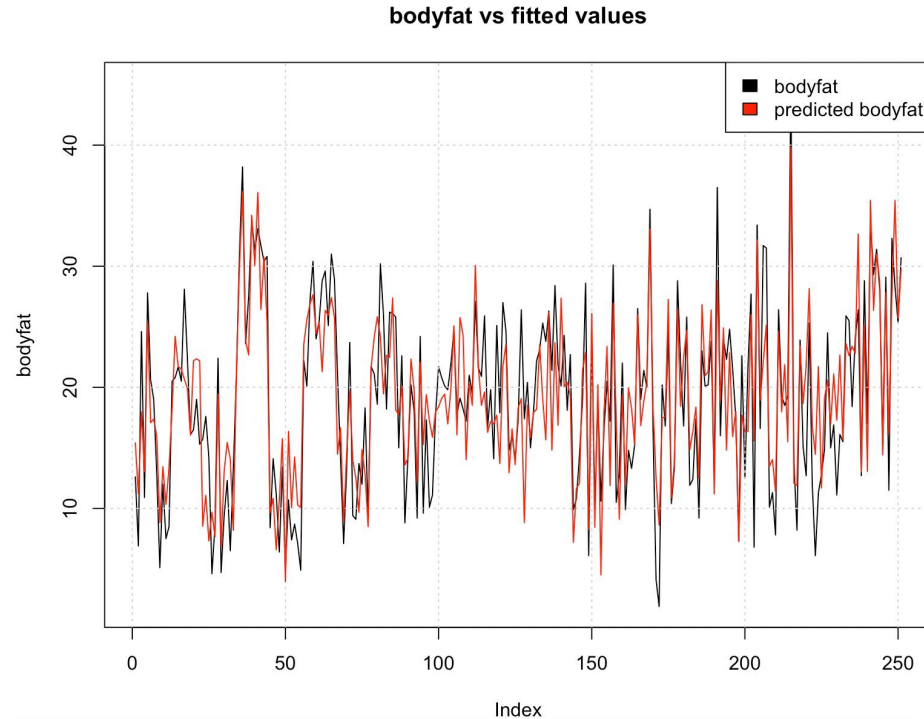
Final Model, Visualization, Model Diagnostics  
, Strengths and Weaknesses

# Final model

$$\text{bodyfat} = -64.4 - 0.008\text{weight} + 1.14\text{abdomen} - 0.003\text{weight} * \text{abdomen}$$

- **Michael Phelps Example: weight(87 kg), abdomen(86.36 cm)**  
Michael's estimated body fat percentage is 8.33%  
This result is pretty close to what we got from Google, which is 10.3%.
- **For A man who has 89 cm abdomen circumference**  
our model predicts that his body fat percentage will decrease 0.275% for every 1 kg increase in his weight.
- **For a man who has 70 kg weight**  
Our model predicts that his body fat percentage will increase 0.93% for every 1 cm increase in his abdomen circumference.

# Visual description of final model



- The trend of fitted values and the original density is roughly the same

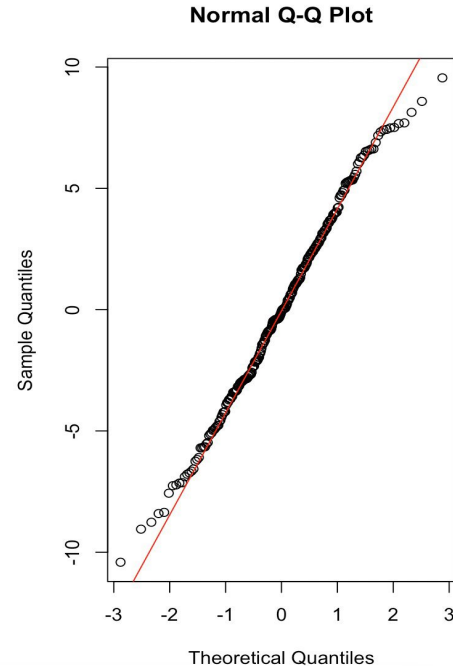
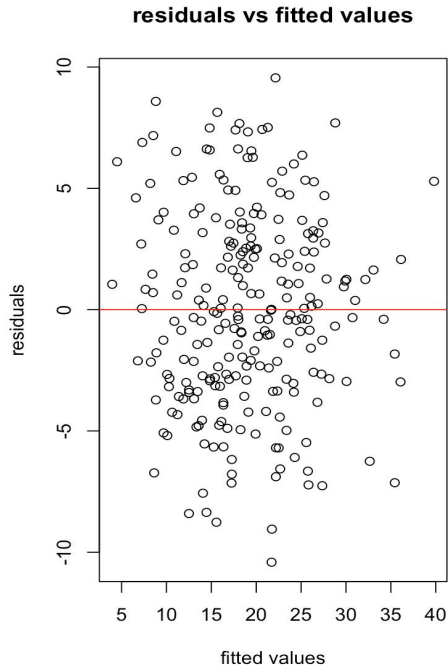
# Statistical properties of final model

<b>predictor</b>	<b>coefficient</b>	<b>p-value</b>	<b>95% lower CI</b>	<b>95% upper CI</b>
<b>intercept</b>	-64.4	4.63e-15	-79.60	-49.25
<b>weight</b>	-0.008	0.94	-0.21	0.19
<b>abdomen</b>	1.14	<2e-16	0.97	1.32
<b>weight*abdomen</b>	-0.003	0.001	-0.005	-0.001

- Abdomen and the interaction term weight\*abdomen are significant under  $\alpha=0.05$
- Overall model is significant with  $p\text{-value}<2e-16$
- $R^2=0.72$ ,  $MSE=16.08$

# Model diagnostics

- We test the normality and randomness of the residuals
- Q-Q plot shows that residuals are approximately normally distributed
- Residuals vs fitted values plot shows that residuals have mean 0 and show no obvious pattern



# Strengths and weaknesses

$$\text{bodyfat} = -64.4 - 0.008\text{weight} + 1.14\text{abdomen} - 0.003\text{weight} * \text{abdomen}$$

- **Strengths**

Simple and Easy to explain

Residuals normally distributed

- **Weaknesses**

Only two predictors

Data is inaccurate for weight and abdomen

Outliers in normal Q-Q plot

# Shiny App

Michael Phelps Example, Extreme Values

# Michael Phelps Example

## Body Fat Calculator

### Personal Information

Note: This body fat calculator is only effective for male since we conduct this study based on a dataset collected among male.

### Weight

Unit

☒ kg  
☐ lb

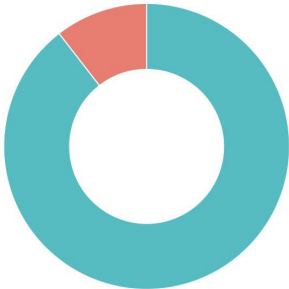
### Circumference of Abdomen

Unit

☒ cm  
☐ inch

Apply Changes

Tissue ■ Fat ■ Other Tissue



Fat

Other Tissue

Your percentage of body fat is: 10.5%.  
Your body fat level is 'Athletes', go on and keep your figure!

Contact us  
E-mail: tjia22@wisc.edu  
E-mail: qzhou85@wisc.edu  
E-mail: khu54@wisc.edu

- We use Michael Phelps's data to introduce our shiny app
- The outcome is 10.5%, which is really close to the Google result 10.3%
- There is some difference between our predicted data and the shiny app outcome due to rounding



# Extreme Values

## Body Fat Calculator

### Personal Information

Note: This body fat calculator is only effective for male since we conduct this study based on a dataset collected among male.

**Weight**

**Unit**

☒ kg  
☐ lb

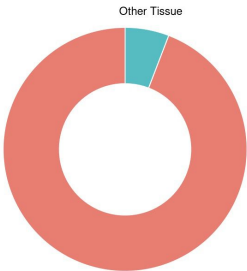
**Circumference of Abdomen**

**Unit**

☒ cm  
☐ inch

[Apply Changes](#)

Tissue Fat Other Tissue



Other Tissue

Fat

Your percentage of body fat is 106.6 %

Oops!! Your body fat is abnormally high. Please check your input.

Contact us  
E-mail: tjia22@wisc.edu  
E-mail: qzhou85@wisc.edu  
E-mail: khu54@wisc.edu

## Body Fat Calculator

### Personal Information

Note: This body fat calculator is only effective for male since we conduct this study based on a dataset collected among male.

**Weight**

**Unit**

☒ kg  
☐ lb

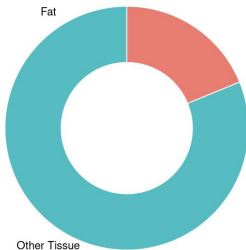
**Circumference of Abdomen**

**Unit**

☒ cm  
☐ inch

[Apply Changes](#)

Tissue Fat Other Tissue



Fat

Other Tissue

Your percentage of body fat is -30.2 %

wOops!! Your body fat is extremely low. Please check your input.

Contact us  
E-mail: tjia22@wisc.edu  
E-mail: qzhou85@wisc.edu  
E-mail: khu54@wisc.edu

# Thank you

Contribution: